PROBABILIDADE E ESTATÍSTICA MATEMÁTICA I

ARA, A.

ara@ufpr.br

Slide 03

DISTRIBUIÇÕES CONDICIONAIS E INDEPENDÊNCIA

DISTRIBUIÇÕES CONDICIONAIS

Anteriormente foram vistas algumas distribuições de probabilidade conjunta para várias variáveis aleatórias (vetor aleatório).

Nesta aula, vamos introduzir as distribuições marginais, condicionais e independência.

DISTRIBUIÇÕES MARGINAIS

Funções de densidade de probabilidade marginal se X e Y são conjuntamente variáveis aleatórias contínuas, então $f_X(\cdot)$ e $f_Y(\cdot)$ são chamadas de funções de densidade de probabilidade marginal.

De maneira geral, seja $X_i, \ldots, X_{im}$ qualquer subconjunto das variáveis aleatórias contínuas conjuntamente $X_1, \ldots, X_k$, $f_{X_i,\ldots,X_{im}}(x_i, \ldots, x_{im}) =$ é chamada de densidade marginal da variável aleatória m-dimensional $(X_i, \ldots, X_{im})$

$$f_X(x) = \int f_{X,Y}(x,y)dy \text{ e } f_Y(x) = \int f_{X,Y}(x,y)dx$$

Analogamente,

$$P(X = x) = \sum_y P(X = x, Y = y) \text{ e } P(Y = y) = \sum_x P(X = x, Y = y)$$

DISTRIBUIÇÕES CONDICIONAIS

Função de densidade condicional discreta

Sejam X e Y variáveis aleatórias conjuntamente discretas com função de probabilidade discreta conjunta $P(X = x, Y = y)$, a função de probabilidade condicional de Y dado $X = x$, denotada por

$$f_{Y|X}(y|x) = \frac{P(X = x, Y = y)}{P(X = x)}$$

A função de probabilidade condicional de Y dado $X = x$, denotada por

$$f_{X|Y}(x|y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

DISTRIBUIÇÕES CONDICIONAIS

Função de distribuição acumulada condicional discreta

Sejam X e Y variáveis aleatórias conjuntamente discretas com função de probabilidade discreta conjunta $P(X = x, Y = y)$, a função de distribuição acumulada condicional discreta de Y dado $X = x$, denotada por

$$F_{Y|X}(y|x) = \sum_{\{j:y_j \leq y\}} f_{Y|X}(y_j|x)$$

Analogamente,

$$F_{X|Y}(x|y) = \sum_{\{j:x_j \leq x\}} f_{X|Y}(x_j|y)$$

DISTRIBUIÇÕES CONDICIONAIS

Função de densidade condicional contínua

Sejam X e Y variáveis aleatórias contínuas distribuídas conjuntamente com função densidade de probabilidade conjunta $f(x, y)$, a função de densidade probabilidade condicional de Y dado $X = x$, denotada por

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f(x)}$$

A função de probabilidade condicional de X dado $Y = x$, denotada por

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f(y)}$$

DISTRIBUIÇÕES CONDICIONAIS

Função de distribuição acumulada condicional contínua

Sejam X e Y variáveis aleatórias contínuas distribuídas conjuntamente com função densidade de probabilidade conjunta $f(x, y)$, a função de distribuição acumulada condicional de Y dado $X = x$, denotada por

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} f_{Y|X}(y|x)dy$$

Analogamente,

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(x|y)dx$$

DISTRIBUIÇÕES CONDICIONAIS

Exemplo: Sejam $X_1, \ldots, X_5$ variáveis aleatórias conjuntamente distribuídas. Seja $W = (X_1, X_2)$ e $V = (X_3, X_5)$

$$f_{W|V}(w, v) = f_{X_1,X_2|X_3,X_5}(x_1, x_2|x_3, x_5) = \frac{f_{X_1,X_2,X_3,X_5}(x_1, x_2, x_3, x_5)}{f_{X_3,X_5}(x_3, x_5)}$$

DISTRIBUIÇÕES CONDICIONAIS

Se X é uma v.a. discreta e Y é qualquer variável aleatória, então $F_{Y|X}(y|x)$ pode ser definida como $P(Y \leq y|X = x)$ com x é um ponto de massa de probabilidade de X.

Se X é uma v.a. contínua, para qualquer evento A, $P(A|X = x)$ não pode ser definido pois $P(X = x|0)$.

Entretanto, se x é tal que os eventos $\{x - h < X < x + h\}$ tem probabilidade positiva para cada $h > 0$, então $P(A|X = x)$ poderia ser definido como

$$P(A|X = x) = \lim_{0<h\to 0} P[A|x - h < X < x + h)$$

desde que o limite exista.

INDEPENDÊNCIA

Independência estocástica: Seja $(X_1, ..., X_k)$ uma variável aleatória k-dimensional, $(X_1, ..., X_k)$ são definidos para serem estocasticamente independentes se e somente se

$$F_{X_1,...,X_k}(x_1, ..., x_k) = \prod_{i=1}^{k} F_{X_i}(x_i)$$

para todos os $x_1, ..., x_k$.

Da mesma forma,

$$f_{X_1,...,X_k}(x_1, ..., x_k) = \prod_{i=1}^{k} f_{X_i}(x_i)$$

OBS: Para X e Y independentes, temos que $f_{X,Y}(x, y) = f_X(x)f_Y(y)$, implica que $f_{X|Y}(x|y) = f_X(x)$ e $f_{Y|X}(y|x) = f_Y(y)$.

ESPERANÇA

Seja $(X_1, \ldots, X_k)$ uma variável aleatória k-dimensional com densidade $f_{X_1,\ldots,X_k}(., \ldots, .)$, o valor esperado de uma função $g(., \ldots, .)$ da variável aleatória k-dimensional é denotado por $E[g(X_1, \ldots, X_k)]$, sendo

- para v. a. k-dimensional discreta:

$$E[g(X_1, \ldots, X_k)] = \sum_{x_1} \ldots \sum_{x_k} g(x_1, \ldots, x_k) f_{X_1,\ldots,X_k}(x_1, \ldots, x_k)$$

- para v. a. k-dimensional contínua:

$$E[g(X_1, \ldots, X_k)] = \int \ldots \int g(x_1, \ldots, x_k) f_{X_1,\ldots,X_k}(x_1, \ldots, x_k) dx_1 \ldots dx_k$$

ESPERANÇA

Teorema: Se $g(x_1, \ldots, x_k) = x_i$ então

$$E[g(x_1, \ldots, x_k)] = E[X_i] = E(X_i) = \mu_{X_i}$$

Denotamos por $\mu_X = (\mu_{X_1}, \ldots, \mu_{X_k})$ o vetor de médias da variável aleatória k-dimensional (vetor aleatório) $X = (X_1, \ldots, X_k)$.

ESPERANÇA

Teorema: Se $g(x_1, \dots, x_k) = (x_i - E(X_i))^2$ então

$$E[g(x_1, \dots, x_k)] = E[(x_i - E(X_i)]^2] = Var(X_i)$$

ESPERANÇA

Exemplo: Suponha $f_{X,Y}(x, y) = (x + y)I_{(0,1)}(x)I_{(0,1)}(y)$

$$E[XY] = \int_0^1 \int_0^1 xy(x + y)dxdy = \frac{1}{3}.$$

$$E[X + Y] = \int_0^1 \int_0^1 (x + y)(x + y)dxdy = \frac{7}{6}.$$

$$E[X] = E[Y] = \frac{7}{12}.$$

ESPERANÇA

Covariância: Seja X e Y duas variáveis aleatórias definidas no mesmo espaço de probabilidade. A covariância de X e Y é definida como

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

desde que a esperança exista.

Resultado: $\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$

ESPERANÇA

Coeficiente de Correlação: O Coeficiente de Correlação de X e Y é definido como

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

desde que a covarância exista, $\sigma_X > 0$ e $\sigma_Y > 0$.

A covariância e o coeficiente de correlação das variáveis aleatórias X e Y são medidas de uma relação linear de X e Y.

ESPERANÇA

Teorema: Se X e Y são v.a. independentes e $g_1(.)$ e $g_2(.)$ duas funções, então

$$E[g_1(X)g_2(Y)] = E[g_1(X)]E[g_2(Y)]$$

Resultado: Se X e Y são v.a. independentes, então $Cov[X, Y] = 0$.

ESPERANÇA

Teorema: Se X e Y são v.a

$$E[g(Y)] = E[E[g(Y)|X]]$$

Particularmente,

$$E[Y] = E[E[Y|X]]$$

Variância condicional: A Variância de Y dado $X = x$ é definida por

$$Var[Y|X = x] = E[Y^2|X = x] - (E[Y|X = x])^2$$

ESPERANÇA

Teorema: $Var[Y] = E[Var[Y|X]] + Var[E[Y|X]]$

Prova:

$$\begin{aligned}
E[Var[Y \mid X]] &= E[E[Y^2 \mid X]] - E[(E[Y \mid X])^2] \\
&= E[Y^2] - (E[Y])^2 - E[(E[Y \mid X])^2] + (E[Y])^2 \\
&= Var[Y] - E[(E[Y \mid X])^2] + (E[E[Y \mid X]])^2 \\
&= Var[Y] - Var[E[Y \mid X]]
\end{aligned}$$

Teorema da Desigualdade de Cauchy-Schwarz: Sejam X e Y com segundos momentos finitos; então $E[XY]^2 \leq E[X^2]E[Y^2]$, com igualdade se e somente se $P[Y = cX] = 1$ para alguma constante c.

Assim, temos que $|\rho_{X,Y}| \leq 1$, com igualdade se e somente se uma variável aleatória é função linear da outra com probabilidade 1.

Teorema: Para as variáveis aleatórias $X_1, \ldots, X_n$

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

e

$$var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} var[X_i] + 2\sum\sum_{i<j} cov[X_i, X_j]$$

OBS: se as variáveis são não correlacionadas, temos que: $var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} var[X_i]$

Particularmente: Teorema do Limite Central

Se $X_1, \ldots, X_n$ são variáveis aleatórias independentes e identicamente distribuídas com média μ_X e variânca σ_X^2, então

$$E(\bar{X}_n) = \mu_X \quad \text{e} \quad \mathrm{Var}(\bar{X}_n) = \frac{\sigma_X^2}{n}$$

com $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$.

A SHORT REVIEW:

- Funções de dist. condicional para variáveis aleatórias discretas
- Funções de dist. condicional para variáveis aleatórias contínuas
- Comentários sobre funções de distribuição condicional
- Independência
- Esperança